Genome-wide identification and characterization of non-specific lipid transfer protein (nsLTP) genes in *Arachis duranensis*

Xiaojun Song
Qingdao Agricultural University
Enguang Li
Qingdao Agricultural University
Hui Song
Qingdao Agricultural University
Guoning Du
Qingdao Agricultural University
Shuai Li
Qingdao Agricultural University
Hong Zhu
Qingdao Agricultural University
Guanxu Chen
Qingdao Agricultural University
Chunmei Zhao
Qingdao Agricultural University
Lixian Qiao
Qingdao Agricultural University
Jingshan Wang
Qingdao Agricultural University
Shanlin Yu
Qingdao Agricultural University
Jiongming Sui (suijiongming@163.com)
Qingdao Agricultural University

Research article

Keywords: Arachis duranensis, nsLTP genes, Abiotic stress, Nematode, Phylogenetic relationship, Co-expression

Posted Date: January 10th, 2020

DOI: https://doi.org/10.21203/rs.2.20551/v1

Abstract

Background: Non-specific lipid transfer proteins (nsLTPs) are known to transfer various lipid molecules between lipid bilayers in plants. In *Arachis duranensis*, however, little is known about the nsLTPs and their responses to biotic and abiotic stresses.

Results: In this study, we identified 64 nsLTP family members (*AdLTPs*) in *A. duranensis*. These *AdLTPs* were classified into six subfamilies (Types 1, 2, C, D, E, and G) and were randomly distributed along nine chromosomes. The K_s value and K_a/K_s value significantly differed between Type 1 and Type D subfamilies. Among paralogs, eight AdLTPs were under positive selection, indicating that these AdLTPs might have different functions in the evolutionary history of *A. duranensis*. qRT-PCR results showed that the expression of *AdLTPs* changed in response to abiotic stresses, including salinity, PEG, low temperature, and ABA. Using RNA-seq data, we also found three *AdLTPs* (*AdLTP1.14*, *AdLTPd8*, and *AdLTPe2*) that were possibly associated with resistance to nematode infection. Among them, *AdLTP1.14*, which belongs to the Type 1 subfamily, was up-regulated at three time points after nematode infection. Co-expression analysis indicated that *DOF* and *WRI1* transcription factors may regulate the *AdLTP* response to nematode infection.

Conclusions: We identified *AdLTPs* in *A. duranensis*. Based on both RNA-seq and qRT-PCR datasets, we found that *AdLTPs* are involved in responses to biotic and abiotic stresses. Our results could provide valuable genomic information for the breeding of peanut cultivars that are resistant to biotic and abiotic stresses.

Background

Non-specific lipid transfer proteins (nsLTPs) are small, basic and abundant proteins in higher plants [1]. nsLTPs contain eight conserved cysteine residues (8CM; C-Xn-C-Xn-CC-Xn-CXC-Xn-C-Xn-C) that form a compact structure stabilized by disulfide bridges. nsLTPs confer resistance to proteolytic enzymes and high temperatures. nsLTP genes are distributed in land plants from the most primitive liverworts and mosses to tracheophytes, but are not found in lower plants such as algae [2]. Plant nsLTPs were previously classified into two subfamilies, including Types 1 and 2 [3]. A subsequent nsLTP classification system retained the two best established subfamilies, Types 1 and 2, and placed the other nsLTP genes in subfamilies, Types C–K [2].

nsLTP genes have different functions [3], including fertilization in flowering plants, adhesion of the pollen tube to the stigma, somatic embryogenesis, lipid metabolism, cutin synthesis, nitrogen fixation, fruit ripening, cell apoptosis, activation and regulation of various signaling cascades, and protection against and adaptation of plants to biotic and abiotic stresses [4–8]. Several nsLTPs have also been found to be major allergens in food and pollen [9–11].

The physiological functions of nsLTP genes in the genus Arachis remain to be clearly elucidated. One member of that genus, the cultivated peanut (A. hypogaea), is an important oil crop in many regions of

the world. The wild ancestral species of the cultivated peanut are generally considered to be A. duranensis and A. ipaensis, which contributed the A and B sub-genomes [12, 13]. Production of the cultivated peanut is often substantially reduced by biotic and abiotic stresses. Interestingly, the stress-resistant capacity is greater in wild peanut than in the cultivated peanut [14, 15]. Stress-resistant genes from wild-type species could therefore be valuable resources for the improvement of the cultivated peanut. The whole genome sequence of the wild ancestral species A. duranensis was recently released [16] and represents an important genomic resource.

In this study, we identified 64 AdLTP genes from A. duranensis. In addition to evaluating the functional diversity of AdLTP genes, we studied their phylogenetic relationships, genome-wide distribution pattern, gene duplication events, and selection pressure. We also used qRT-PCR to identify AdLTP genes that might contribute to abiotic stress, and we used RNA-seq data previously reported for nematode-infected plants [17] to identify AdLTP genes that potentially confer resistance to nematode infection. Our results provide a comprehensive analysis of AdLTP genes and provide preliminary information on specific AdLTP genes that may be involved in biotic and abiotic stress resistance.

Results

Identification and characterization of nsLTP genes (AdLTPs) in A. duranensis

Using a bioinformatics approach, we identified a total of 64 AdLTP genes (Tables S1). The structure and characteristics of AdLTPs were analyzed based on the genomic sequences, coding domain sequences (CDS), and amino acid sequences (Table 1). The genomic lengths of the AdLTPs ranged from 276 bp to 4572 bp, CDS lengths ranged from 276 bp to 735 bp, and the deduced number of amino acids ranged from 91 to 244. The molecular weights ranged from 9278.82.74 to 26135.12 Da, and the isoelectric points ranged from 4.09 to 10.10 (Table 1).

Table 1
nsLTP genes (AdLTPs) in the A. duranensis genome.

Gene ID	Chromosome	Genomic length (bp)	CDS length (bp)	No. of AA	Mol. Wt. (Da)	pI
AdLTP1.1	A07	1603	360	119	13280.79	8.79
AdLTP1.2	A08	443	330	109	12725.74	5.64
AdLTP1.3	A07	442	348	115	11618.49	8.79
AdLTP1.4	A06	974	384	127	13374.64	8.69
AdLTP1.5	A04	821	363	120	12517.84	8.85
AdLTP1.6	A04	1761	432	143	15716.65	8.82
AdLTP1.7	A02	1819	357	118	11955.19	9.17
AdLTP1.8	A02	1080	351	116	11760.82	9.04
AdLTP1.9	A02	333	333	110	12159.45	8.91
AdLTP1.10	A02	724	381	126	13959.97	10.10
AdLTP1.11	A02	1181	375	124	13451.19	9.50
AdLTP1.12	A09	357	357	118	12422.54	4.09
AdLTP1.13	A02	1229	351	116	11539.7	9.28
AdLTP1.14	A01	1242	420	139	15227.89	8.96
AdLTP1.15	A02	906	354	117	11777.91	9.24
AdLTP2.1	A05	6138	588	195	20553.99	8.76
AdLTP2.2	A04	285	285	94	9783.83	8.94
AdLTP2.3	A05	339	339	112	11759.73	8.67
AdLTPc1	A05	276	276	91	9278.82	8.16
AdLTPd1	A01	312	312	103	11005.21	8.47
AdLTPd2	A03	324	324	107	11578.77	8.59
AdLTPd3	A07	309	309	102	10487.56	9.22
AdLTPd4	A09	321	321	106	11218.63	6.78
AdLTPd5-1	A09	315	315	104	10882.09	8.41
AdLTPd5-2	A09	315	315	104	11057.32	8.63

Gene ID	Chrom osome	Genomic length (bp)	CDS length (bp)	No. of AA	Mol. Wt. (Da)	pI
AdLTPd5-3	A09	315	315	104	10884.06	8.42
AdLTPd5-4	A09	291	291	96	9984.97	5.98
AdLTPd5-5	A09	291	291	96	9984.97	5.98
AdLTPd6	A01	1485	366	121	12755.20	10.09
AdLTPd7	A02	1048	360	119	12581.69	8.66
AdLTPd8	A03	1082	342	113	11923.14	8.39
AdLTPe1	A08	396	396	131	13176.65	8.06
AdLTPe2	A08	547	327	108	11632.81	7.49
AdLTPg1	A05	827	663	220	22389.62	8.11
AdLTPg2	A02	1713	459	152	15716.36	4.86
AdLTPg3	A05	2342	495	164	17781.55	4.77
AdLTPg4	A04	4572	603	200	21540.10	6.92
AdLTPg5	A09	819	492	163	16162.52	8.43
AdLTPg6	A01	2413	585	194	20011.10	8.94
AdLTPg7	A05	1808	453	150	16267.04	8.07
AdLTPg8	A01	1713	426	141	14844.29	7.53
AdLTPg9	A04	1318	534	177	19516.96	8.61
AdLTPg10	A05	886	735	244	26135.12	8.78
AdLTPg11	A03	918	585	194	20636.9	7.5
AdLTPg12	A04	3591	603	200	19805.61	4.68
AdLTPg13	A08	2768	552	183	19483.32	8.38
AdLTPg14	A03	744	591	196	20491.68	8.44
AdLTPg15	A06	480	396	131	13584.81	6.5
AdLTPg16	A02	922	597	198	19585	7.48
AdLTPg17	A06	2481	555	184	19100.78	8.63
AdLTPg18	A03	1360	591	196	19254.96	8.06
AdLTPg19	A01	522	522	173	17003.52	9.08

Gene ID	Chromosome	Genomic length (bp)	CDS length (bp)	No. of AA	Mol. Wt. (Da)	pI
AdLTPg20	A01	450	450	149	15059.6	9.13
AdLTPg21	A05	977	639	212	21138.36	8.09
AdLTPg22	A01	396	396	131	13750.4	8.93
AdLTPg23	A05	828	468	155	15809.22	8.96
AdLTPg24	A01	519	519	172	18281.06	8.86
AdLTPg25	A01	414	414	137	14497.91	6.01
AdLTPg26	A06	372	372	123	12535.15	8.57
AdLTPg27	A01	441	441	146	15427.99	8.07
AdLTPg28	A04	387	387	128	13344.99	8.59
AdLTPx1	A08	483	483	160	18757.85	5.96
AdLTPx2	A08	438	438	145	16921.19	5.73
AdLTPx3	A08	495	495	164	19338.48	8.16

we used the gene structure display server program to investigate AdLTP exon–intron organization. The results revealed that 13 of the 15 AdLTP1 members and 3 of the 12 AdLTPd members had the 2-exon conserved gene structure, while the other genes contained only one exon in these two subfamilies. The members of subfamily 2 and subfamily G contained 1–3 exons and 1–4 exons, respectively (Fig. 1).

Classification and phylogenetic analysis of AdLTPs

To ensure accurate inference of the topological structures, we used computationally efficient maximum likelihood and Bayesian inference method to construct phylogenetic trees. Based on Edstama et al. [2], nsLTP proteins can be classified into six subfamilies. In A. duranensis, we found 15, 3, 1, 12, 2, and 28 genes from subfamily 1, 2, C, D, E, and G, respectively (Fig. 2). In addition, 3 nsLTPs (AdLTPx1, AdLTPx2, and AdLTPx3) were singles. The phylogeny showed that Types1, 2, C, and E formed monophyletic groups. Types D and G, however, did not form supported monophyletic groups, i.e., they were interspersed among other monophyletic groups. The spacing patterns of most Cys residues in the six classified AdLTPs (Table 2) agreed with those reported by Edstam et al. for other species [2], but the number of bases between the cysteins in a few AdLTPs was greater than that reported by Edstam et al. (Additional file 2).

Table 2
Spacing pattern of the six types of AdLTPs in A. duranensis.

Type	Spacing pattern									
1	C	9	C	13–16	CC	17–21	C-1-C	21–24	C	4–13
2	C	7	C	13	CC	8	C-1-C	23	C	6–41
C	C	9	C	16	CC	9	C-1-C	12	C	3
D	C	9–15	C	14–19	CC	9–12	C-1-C	22,24	C	6–10
E	C	13	C	15	CC	9	C-1-C	22	C	6

Chromosomal locations and gene duplication events of AdLTPs

AdLTP genes were mapped to their chromosome positions using the A. duranensis genome database. AdLTP genes were distributed on 9 of the 10 chromosomes in the genome (Fig. 2). Chromosome A01 contained the largest number of AdLTP genes (11), and chromosome A10 did not contain any AdLTP genes.

Gene duplication often occurs during gene evolution [18]. In this study, we found one interchromosomal duplication (AdLTPg19/AdLTPg20). An approximately 43-kb region on chromosome A09 included the following six AdLTPds (Fig. 1): AdLTPd4, AdLTPd5-1, AdLTPd5-2, AdLTPd5-3, AdLTPd5-4, and AdLTPd5-5. This region also contained two tandems (AdLTPd5-1,/AdLTPd5-2 and AdLTPd5-3/AdLTPd5-4) and one segmental duplication (AdLTPd5-4/AdLTPd5-5) (Fig. 3). These results indicate that gene duplications have been a driving force in AdLTP evolution.

Multiple sequence alignments showed that AdLTP proteins contained a typical conserved C-Xn-C-Xn-CC-Xn-CXC-Xn-C-Xn-C domain (where C is cysteine, and X is any other amino acid) (Fig. 4). The results indicate that these nsLTP genes might have a similar biological function and a common origin.

To assess the selection pressure on AdLTPs, we used the CODEML of the PAML package in order to compute the K_a, K_s, and ω values. We found that K_a did not significantly differ ($P > 0.05$) among Types 1, 2, D, and G. However, K_s significantly differed ($P = 0.033$) between Types 1 and D, and ω significantly differed ($P = 0.043$) between Types 1 and D. Although the average ω value among the AdLTP genes was < 1, 8 genes were found to be under positive selection; these included three genes in Type 1, one gene in Type D, and four genes in Type G (Fig. 5). These results suggest functional diversity among those AdLTP genes that underwent positive selection.

Expression of AdLTP genes in 22 tissues

Although the PeanutBase website contains RNA-seq datasets from 22 A. duranensis tissues, only 44 AdLTP genes were expressed (Fig. 6). Based on their expression in the 22 tissues, the 44 genes were classified into clades I, II, and III. Among the 44 genes, expression levels were high for 12 genes in clade I, were low for 18 genes in clade III, and were intermediate for 14 genes in clade II (Fig. 6). Most genes in the LTP1 and LTPe subfamilies were expressed in aboveground organs, while most genes in the LTPd and LTPg subfamilies were expressed in belowground organs and especially in seeds (in the case of LTPds) and in roots (in the case of LTPgs) (Fig. 6).

Expression of AdLTP genes in response to abiotic stress and nematode infection

To increase our understanding of AdLTP regulation in response to abiotic stress, we subjected A. duranensis seedlings to stresses induced by PEG, low temperature, salinity, and ABA treatment. We then used real-time PCR (qRT-PCR) to assess the transcription levels of 45 AdLTP genes (8 AdLTP1s, 3 AdLTP2s, 1 AdLTPcs, 4 AdLTPds, 2 AdLTPes, 24 AdLTPgs, and 3 AdLTPxs). The results showed that the expression of all 45 AdLTP genes was modified in response to at least one of the stresses (Fig. 7). Some members of the same subfamily had similar expression patterns; this was true for 7 AdLTPgs in response to PEG treatment (AdLTPg7, AdLTPg9, AdLTPg13, AdLTPg15, AdLTPg19, AdLTPg24, and AdLTPg28, Fig. 7a), 3 AdLTPds in response to low temperature treatment (AdLTPd2, AdLTPd3, and AdLTPd6, Fig. 7b), 3 AdLTP1s in response to salinity treatment (AdLTP1.3, AdLTP1.4, and AdLTP1.15, Fig. 7c), and 5 AdLTPgs in response to ABA treatment (AdLTPg7, AdLTPg14, AdLTPg17, AdLTPg22, and AdLTPg23, Fig. 7d). Most members of the same subfamily, however, had different expression patterns in response to a particular stress (Fig. 7).

The RNA-seq results based on the PeanutBase website revealed that three AdLTP genes (AdLTP1.14, AdLTPd8, and AdLTPe2) were differentially expressed in root tissues after 3, 6, and 9 days of nematode infection, indicating that these genes are possibly involved in responses to nematode infection. Among them, the expression level of AdLTP1.14 was up-regulated at three time points. Interesting, AdLTPe2 and AdLTP1.14 showed the highest expression level in roots and nodule roots, respectively.

In a previous study of nematode-infected A. duranensis plants, co-expression analysis was used to assign 462 genes A. duranensis to five modules [19]. AdLTP1.14, AdLTPd8, and AdLTPe2 were in module 5. Analysis of the cis-acting elements showed that the sequences upstream of these three nsLTP genes can bind to certain transcription factors including Dof, WRI1, GT-1, HSF, and TFHP-1 (Additional file 3). The co-expression analysis also indicated that the Dof and WRI1 (AP2 family) transcription factors may regulate the response of AdLTP1.14, AdLTPd8, and AdLTPe2 to nematode infection. In addition, we found other A. duranensis genes that may interact with AdLTP genes and have potential roles in resisting nematode infection or biotic stress based on previous studies (Additional file 4). When inoculated with nematodes, for example, transgenic soybean plants overexpressing the expansin-like B gene from peanut had a substantially reduced number of galls [17]. Leucine-rich repeat (LRR)-containing genes and trypsin inhibitors have been found to be important in plant resistance to biotic stress [19, 20]. Our results indicated that AdLTP1.14, AdLTPd8, and AdLTPe2 may interact with expansin, LRR-containing protein,

and trypsin inhibitors (Additional file 4). Furthermore, three groups of genes (AdLTPe2, Aradu.7623G, Aradu.MKB35, and Aradu.Y057X; AdLTPd8 and Aradu.CG0MN; and AdLTP1.14, Aradu.B5LWV, and Aradu.UIP3J) had similar expression patterns, respectively (Fig. 8). These results indicated that these genes may have synergistic effects in resisting nematode infection.

Discussion

In the current study, we investigated the nsLTP genes from A. duranensis because such genes may help explain the substantial resistance of this non-cultivated species to abiotic and biotic stresses [14, 15] and because the acquired information could be useful for breeding stress-resistant cultivars of A. hypogaea (the cultivated peanut). We identified and characterized 64 nsLTP genes from A. duranensis and assigned these AdLTPs to six subfamilies: Types 1, 2, C, D, E, and G.

Having identified and characterized 64 AdLTP genes, we next investigated duplication of AdLTPs. Gene duplication is common in plants and provides resources for novel gene functions. Some duplicated genes become pseudogenized and have no function, while other gene duplication pairs evolve new functions [21, 22]. We identified one interchromosomal duplication (AdLTPg19/ AdLTPg20), two tandem duplications (AdLTPd5-1/AdLTPd5-2 and AdLTPd5-3/AdLTPd5-4), and one segmental duplication (AdLTPd5-4/AdLTPd5-5) in A. duranensis (Fig. 2). These five AdLTPd5 genes were distributed in a 43-kb region on chromosome A09 (Fig. 2). Some duplicated gene pairs showed similar expression patterns; for example, AdLTPd5-1, AdLTPd5-2, AdLTPd5-3, AdLTPd5-4, and AdLTPd5-5 had similar expression pattern with higher expression in seeds and lower expression in most aboveground organs, such as leaves and flowers, suggesting that they may function in seed growth and development. Interestingly, AdLTPd5-1/AdLTPd5-2 and AdLTPd5-4/AdLTPd5-5 showed identical expression levels in 22 A. duranensis tissues, suggesting that the functions of these gene pairs have been conserved. The expression of other duplicated genes, in contrast, differed between the gene pairs across tissues or growth phases. AdLTPd5-4, for example, had higher expression levels in seeds, pods, and stamens than AdLTPd5-3, suggesting a functional divergence for these two duplicated AdLTPd genes (Fig. 6).

Our qRT-PCR results showed that the AdLTP genes were induced by at least one abiotic stress treatment and that most genes were induced by all four abiotic stress treatments including PEG, NaCl, low temperature, and ABA. One possible explanation for why AdLTP genes are induced by abiotic stress is that the proteins encoded by these genes may be involved in the biosynthesis of the cuticle layer. The cuticle has been found to be important in maintaining water balance under a variety of stresses, and the induction of nsLTPs in response to stress is accompanied by a thickening of the cuticle [23] and a disruption of the glycosylphosphatidylinositol-anchored LTPg gene, which leads to changes in the lipid composition and density of the cuticle [24]. Another possible explanation for why AdLTP genes are induced by abiotic stress is that nsLTP proteins can bind to lipids; by binding to lipids in thylakoid membranes, nsLTP proteins stabilize the membranes under stress conditions [25, 26].

nsLTPs belong to a large family of pathogenesis-related proteins (PRPs). Induction of the synthesis of these proteins occurs upon exposure of a plant to pathogens and underlies a key plant defense mechanism [27]. Peanut yields are reduced by pathogenic fungi, bacteria, viruses, and nematodes [28]. Peanuts infected by the root knot nematode, Meloidogyne arenaria, become stunted and wilt, and have enhanced susceptibility to other pathogens [29]. In the current study, we found that three AdLTP genes were distributed in the same co-expression module and have potential roles in resisting nematode infection. The co-expression analysis indicated that some transcription factors may regulate the response of nsLTP genes to nematode infection. Among these transcription factors, WRI1 is known to affect cutin synthesis[30], and Dof transcription factors may help resist nematode infection [19]. We also found that AdLTP proteins may interact with expansin, LRR protein, and trypsin inhibitor. Additional study of AdLTP genes could clarify the gene regulatory networks that increase resistance to nematode infections.

The protective function of LTPs in plants is related to their antimicrobial and cryoprotective activities, their ability to inhibit exogenous enzymes, and their possible involvement in the secretion of other components of the plant immune system. For example, some members of the LTP class, like protease inhibitors (PRP-6) and certain defensins (PRP-12), can inhibit the activity of proteolytic enzymes and α-amylases [31, 32]. Barley seed LTPs of both subclasses were found to inhibit cysteine endoproteases [33]. LTP1 from Ginkgo biloba seeds inhibits cysteine (papain), aspartate (pepsin), and serine (trypsin) proteases [34]. LTP1 from coffee and pepper seeds inhibits the activity of human α-amylase [35, 36].

We found that four allergens from Arachis (Ara h2 [LTPx1], Ara h6 [LTPx2], Ara h7 [LTPx3], and Ara h9 [LTP1.13], were nsLTP proteins, and all four were responsive to the four abiotic stress treatments (Fig. 7). Moreover, LTPx1, LTPx2, and LTPx3 were found to be located on a 90-kb segment of chromosome A08 and are α-amylase inhibitors according to GenBank (http://www.ncbi.nlm.nih.gov/ genbank). Orthologous genes of AdLTPx2 and AdLTP1.13, however, have not been found in A. ipaensis, which indicates that different wild ancestral species of the cultivated peanut may supply unique genes.

Conclusions

We identified the number and type of nsLTP genes in A. duranensis. We further estimated the substitution rate of each type of nsLTP domain between paralogs. nsLTP domains were inferred to have mainly been subjected to positive selection. In addition, we comprehensively identified the nsLTP genes that were involved in responses to biotic and abiotic stresses. We also found that WRI1 and Dof transcription factors may regulate the nsLTP genes associated with resistance to nematode infection. Our results may be useful for breeding peanut cultivars with increased resistance to abiotic and biotic stresses.

Methods

Identification and characterization of nsLTP genes in A. duranensis

The genome sequence of A. duranensis (Aradu.V14167.a1.M1) was downloaded from http://peanutbase.org [16]. The Hidden Markov Model (HMM) profile of the nsLTP domain (PF00234 and PF14368) was downloaded from the pfam database (http://pfam.janelia.org). The Arabidopsis thaliana nsLTP protein (AtLTP) sequences were downloaded from the Arabidopsis Information Resource website (TAIR, http://www.arabidopsis.org). The predicted nsLTP proteins of A. duranensis and Arabidopsis were extracted using the HMMER program with default parameters. To verify the reliability of results, all protein sequences were checked in the pfam database.

Phylogenetic and conservation analysis of A. duranensis nsLTP proteins

Multiple sequence alignments of AtLTP and AdLTP proteins were performed with the L-INS-I method in MAFFT 7.0 [36]. AtLTP proteins were used as query to categorize the AdLTP proteins based on the phylogenetic tree. Phylogenetic analyses were performed using maximum likelihood (ML) and Bayesian inference (BI) methods. ML was performed in PhyML 3.1 [38] with 100 bootstrap replicates. BI was computed in MrBayes v3.2.6 [39] with a mixed amino acid substitution model. Posterior probabilities were calculated over 10,000,000 generations via two independent runs of four simultaneous Markov chain Monte Carlo chains with every 5000th tree saved. The conservation of different nsLTP subfamily proteins was detected by weblogo online software (http://weblogo.berkeley.edu/logo.cgi). Highly variable parts of the alignment were determined and manually excluded in SeaView v4 [40].

Genome-wide distribution patterns, gene duplication events, and selection pressure of AdLTP genes

Information on the chromosomal location of AdLTP proteins was obtained from PeanutBase (http://peanutbase.org). AdLTP genes were exhibited on different chromosomes using MapInspect software (http://mapinspect.software.informer.com/). We defined AdLTP gene duplication events based on the following criteria: (1) length of aligned sequences > 80% of each sequence length; (2) identity ≥ 70%; and (3) E-value ≤ 10E − 10. Non-synonymous (K_a) and synonymous (K_s) substitution values of subfamilies of AdLTP genes were calculated by the CODEML program in the PAML software package [41]. The nonsynonymous/synonymous substitution rate ratio (ω = Ka/Ks) can generally indicate the nature of selective pressures, i.e., ω = 1, < 1, and > 1 indicate neutral evolution, purifying selection, and positive selection, respectively, on the target gene.

Expression of AdLTP genes in different tissues and after nematode infection

RNA-seq datasets from 22 A. duranensis tissues have been previously made available on the PeanutBase website (https://peanutbase.org/gene_expression/atlas) [42, 43]. The differentially expressed genes in root tissue after 3, 6, and 9 days of nematode infection have also been published and are available on PeanutBase (https://peanutbase.org/gene_expression/atlas_nematode). In the current study, we generated heat maps for these sequences using the heatmap script of the R program in the gplots CRAN library package. The fragments per kilobase of transcript per million mapped reads (FPKMs) value for each gene were normalized using a $\log_2$ (fold-change) value.

Plant materials and stress treatments

Seeds of wild A. duranensis PI219823 were planted in pots containing sandy soil and were kept at 28℃ with 3000 lux light intensity and an 11-h dark/13-h light cycle. For stress treatments, 6-week-old seedlings were irrigated with 250 mM NaCl, 20% PEG, or 100 mM ABA, or were kept at 4℃. The leaves were collected at 0, 6, 12, 24, and 48 h after stress treatment, and were immediately frozen in liquid nitrogen for RNA extraction or were stored at -80℃ until used.

Gene expression analysis by qRT-PCR

Total RNA was extracted using the CTAB method [44]. The actin gene was used as a reference gene to quantify AdLTP gene expression [45]. All primers used in this study are listed in Additional file 1. qRT-PCR reactions were carried out using fluorescent dye SYBR-Green (Takara, Dalian, China) and Fast Start Universal SYBR Green Master (ROX) with a 7500 real-time PCR machine (ABI). The reactions were carried out as follows: 30 s at 95℃ for denaturation, followed by 40 cycles of 5 s at 95℃ and 30 s at 60℃. Three biological replicates were used. Data were quantified using the $2^{-\Delta\Delta Ct}$ method based on Ct values of AdLTP and actin genes. Values were considered to be significant at $P < 0.05$.

Co-expression analyses

Using previously published data and a weighted gene co-expression network analysis (WGCNA) script in R [46], we identified genes that were co-expressed in 22 tissues of A. duranensis in response to nematode infection. Differentially expressed genes with a $|\log_2$ (fold-change)| value ≧ 2 were used for WGCNA analyses. A soft threshold (β) value of 12 was used in the transformation of the adjacency matrix in order to meet the scale-free topology criteria. Co-expression modules were created with the blockwise module function according to the method introduced by Song [19]. Gene ontology (GO) annotations for genes in each module containing the related gene were extracted from the A. duranensis genome available on the PeanutBase website (https://peanutbase.org/files/genomes/Arachis_ duranensis/) [42].

Abbreviations

nsLTP, non-specific lipid transfer protein; qRT-PCR, quantitative real-time PCR; CDS, coding domain sequence; RNA-seq, RNA-sequencing; LRR, Leucine-rich repeat; WRI1, WRINKLED1; PRP, pathogenesis-related protein; Dof, DNA binding with one finger; HSF, heat shock transcription factor; WGCNA, co-expression network analysis; ML, maximum likelihood; BI, Bayesian inference; HMM, Hidden Markov Model; GO, Gene ontology; FPKM, fragments per kilobase of transcript per million mapped reads; A. hypogaea, Arachis hypogaea; A. duranensis, Arachis duranensis; A. ipaensis, Arachis ipaensis; At, Arabidopsis thaliana

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by the National Natural Science Foundation of China (31571705 and 31872875) and The Cooperation and Innovation of Peanut Industry in Shandong Province. The funders had no role in the design of the study, collection, analysis, and interpretation of data, or in writing the manuscript.

Authors' contributions

JS conceived and designed the research. XS, NL, HS, GD, SL, HZ, CZ, LQ, JW and SYconducted the experiments and analyzed the data. JS wrote the manuscript. All authors read and approved the manuscript.

Acknowledgements

We thank Xiaojun Zhang at Qingdao Agricultural University for supplying wild peanut seeds PI219823.

References

1. Shewry PR, Beaudoin F, Jenkins J, Griffifiths-Jones S, Mills ENC: Plant protein families and their relationships to food allergy. Biochem. Soc. Transact. 2002; 30:906-910.
2. Edstam MM, Viitanen L, Salminen TA, Edqvist J: Evolutionary history of the non-specifc lipid transfer proteins. Mol Plant, 2011; 4(6): 947-964.
3. Boutrot F, Chantret N, Gautier MF: Genome-wide analysis of the rice and *Arabidopsis* non-specific lipid transfer protein (nsLtp) gene families and identification of wheat nsLtp genes by EST data mining. BMC Genomics. 2008; 9:86.
4. Park SY, Jauh GY, Mollet JC, Eckard KJ, Nothnagel EA, Walling LL, Lord EM: A lipid transfer-like protein is necessary for lily pollen tube adhesion to an in vitron matrix. Plant Cell. 2000; 12:151-163.

5. Tomassen, MM, Barrett DM, van der Valk HC, Woltering EJ: Isolation and characterization of a tomato non-specifific lipid transfer protein involved in polygalacturonase- mediated pectin degradation. J. Exp. Bot. 2007; 58:1151-1160.
6. Lascombe MB, Bakan B, Buhot N, Marion D, Blein JP, Larue V, Lamb C, Prangé T: The structure of defective in induced resistance protein of *Arabidopsis thaliana*, DIR1, reveals a new type of lipid transfer protein. Protein Sci. 2008; (17)9:1522-1530.
7. DeBono A, Yeats TH, Rose JKC, Bird D, Jetter R, Kunst L, Samuels L: *Arabidopsis* LTPG is a glycosylphosphatidylinositol-anchored lipid transfer protein required for export of lipids to the plant surface Plant Cell. 2009; 2l(4): 1230-1238.
8. Pii Y, Astegno A, Peroni E, Zaccardelli M, Pandolfifini T, Crim, M: The *Medicago truncatula* N5 gene encoding a root-specifific lipid transfer protein is required for the symbiotic interaction with *Sinorhizobium meliloti*. Mol. Plant Microbe Interact. 2009; 22: 1577-1587.
9. Salcedo G, Sanchez-Monge R, Diaz-Perales A, Garcia-Casado G, Barber D: Plant non-specifific lipid transfer proteins as food and pollen allergens. Clin. Exp. Allergy. 2004; 34: 1336-1341.
10. Lauer I, Dueringer N, Pokoj S, Rehm S, Zoccatelli G, Reese G, Miguel-Moncin MS, Cistero-Bahima A, Enrique E, Lidholm J, Vieths S, Scheurer S: "The non-specific lipid transfer protein, Ara h 9, is an important allergen in peanut," Clinical and Experimental Allergy. 2009; (39)9:1427-1437.
11. Zoccatelli G, Pokoj S, Foetisch K, Bartra J, Valero A, del Mar San Miguel-Moncin M, Vieths S, Langfelder P, Horvath S: WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559.
12. Seijo J, Lavia G, Fernandez A, krapovickas A, Ducasse D, Moscone E. Physical mapping of the 5S and 18S-25S rRNA genes by FISH as evidence that *Arachis duranensis* and *ipaënsis* are the wild diploid progenitors of *A. hypogaea* (*Leguminosae)*. Am J Bot. 2004; 91:1294-303.
13. Ramos ML, Fleming G, Chu Y, Akiyama Y, Gallo M, Ozias-Akins P: Chromosomal and phylogenetic context for conglutin genes in *Arachis* based on genomic sequence. Mol Genet Genomics. 2006; 275: 578-592.
14. Pande S, Narayana Rao J: Resistance of wild Arachis species to late leaf spot and rust in greenhouse trials. Plant Dis. 2001; 85:851-855.
15. Simpson CE: Use of wild Arachis species/introgression of genes into A. hypogaea L. Peanut Sci. 2001; 28:114-116.
16. Bertioli DJ, Cannon SB, Froenicke L, Huang G, Farmer AD, Cannon EKS, Liu X, Gao D, Clevenger J, Dash S, Ren L, et al: The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. Nat Genet. 2016; 48:438-46.
17. Guimaraes PM, Guimaraes LA, Morgante CV, Silva OB Jr, Araujo AC, Martins AC, Saraiva MA, Oliveira TN, Togawa RC, Leal-Bertioli SC, Bertioli DJ, Brasileiro AC: Root transcriptome analysis of wild peanut reveals candidate genes for nematode resistance. PLoS ONE. 2015; 10(10): e0140937.
18. Adams KL, Wendel JF: Polyploidy and genome evolution in plants. Curr Opin Plant Biol. 2005; 8(2):135- 141.

19. Song H, Guo ZL, Chen T, Sun J, Yang GF: Genome-wide identification of LRR-containing sequences and the response of sequences to nematode infection in *Arachis duranensis*. BMC Plant Biology. 2018; 18:279.
20. Zhao H, Ullah H, McNeill MR, Du G, Hao K, Tu X, Zhang Z: Inhibitory Effects of Plant Trypsin Inhibitors *Msti-94*and*Msti-16* on *Therioaphis trifolii* (Monell) (Homoptera: Aphididae) in Alfalfa. Insects. 2019; 10(6): 154.
21. Kondrashov FA, Rogozin IB, Wolf YI, Koonin EV: Selection in the evolution of gene duplications. Genome Biol. 2002; 3(2):0008
22. Wang Z, Zhou Z, Liu Y, Liu T, Li Q, Ji Y, Li C, Fang C, Wang M, Wu M, Shen Y, Tang T, Ma J, Tian Z: Functional evolution of phosphatidylethanolamine binding proteins in soybean and *Arabidopsis*. Plant Cell. 2015; 27(2):323-336.
23. Cameron KD, Teece MA, Smart LB: Increased accumulation of cuticular wax and expression of lipid transfer protein in response to periodic drying events in leaves of tree tobacco. Plant Physiol. 2006; 140: 176-183.
24. Lee SB, Go YS, Bae HJ, Park JH, Cho SH, Cho HJ, Lee DS, Park OK, Hwang I, Suh MC: Disruption of glycosylphosphatidylinositol-anchored lipid transfer protein gene altered cuticular lipid composition, increased plastoglobules, and enhanced susceptibility to infection by the fungal pathogen *Alternaria brassicicola*. Plant Physiol. 2009; 150(1): 42-54.
25. Hincha DK: Cryoprotectin: a plant lipid-transfer protein homologue that stabilizes membranes during freezing. Philos Trans R Soc Lond B Biol Sci. 2002; 357(1423):909-916.
26. Sror HA, Tischendorf G, Sieg F, Schmitt JM, Hincha DK: Cryoprotectin protects thylakoids during a freeze-thaw cycle by a mechanism involving stable membrane binding. Cryobiology. 2003; 47(3):191-203.
27. Hoffmann-Sommergruber K: Pathogenesis-related (PR)-proteins identified as allergens. Biochem Society Transactions. 2002; 30:930-935.
28. Burow MD, Starr JL, Park CH, Simpson CE, Paterson AH: Introgression of homeologous quantitative trait loci (QTLs) for resistance to the root-knot nematode [*Meloidogyne arenaria* (Neal) Chitwood] in an advanced backcross-QTL population of peanut (*Arachis hypogaea* L.). Mol Breeding. 2014; 34(2):393-406.
29. Tirumalaraju SV, Jain M, Gallo M: Differential gene expression in roots of nematode-resistant and -susceptible peanut (*Arachis hypogaea*) cultivars in response to early stages of peanut root-knot nematode (*Meloidogyne arenaria*) parasitization. J Plant Physiol. 2011;168: 481-92.
30. Hao SX, Ma YY, Zhao S, Ji QL, Zhang KZ, Yang MF, Yao YC (). McWRI1, a transcription factor of the AP2/SHEN family, regulates the biosynthesis of the cuticular waxes on the apple fruit surface under low temperature. PLoS One. 2017; 12 (10): e0186996.
31. Zhang N, Jonnes BL, Tao HP: Purification and characterization of a new class of insect α-amylase inhibitors from barley. Cereal Chem. 1997; (74)2: 119-122.

32. Melo FR, Rigden DJ, Franco OL, Mello LV, Ary MB, Grossi de Sa MF, Bloch C: Inhibition of trypsin by cowpea thionin: Characterization, molecular modeling, and docking. Proteins. 2002; (48)2: 311-319.
33. Jones BL, Marinac LA: Purification and partial characterization of a second cysteine proteinase inhibitor from ungerminated barley (*Hordeum vulgare* L.). J Agric Food Chem. 2000; 48:257-264.
34. Sawano Y, Hatano K, Miyakawa T, Komagata H, Miyauchi Y, Yamazaki H, Tanokura M: Proteinase inhibitor from ginkgo seeds is a member of the plant nonspecific lipid transfer protein gene family. Plant Physiol. 2008; (14) 4: 1909-1919.
35. Diz MS, Carvalho AO, Ribeiro SF, Da Cunha M, Beltramini L, Rodrigues R, Nascimento VV, Machado OL, Gomes VM: Characterisation, immunolocalisation and antifungal activity of a lipid transfer protein from chili pepper (*Capsicum annuum*) seeds with novel α-amylase inhibitory properties. Physiol Plant. 2011; (14) 3: 233-246.
36. Zottich U, Da Cunha M, Carvalho AO, Dias GB, Silva NC, Santos IS, do Nacimento VV, Miguel EC, Machado OL, Gomes VM: Purification, biochemical characterization and antifungal activity of a new lipid transfer protein (LTP) from Coffea canephora seeds with α-amylase inhibitor properties. Biochim Biophys Acta. 2011; (1810)4: 375-383
37. Katoh K, and Standley DM: MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Bio. Evol. 2013; 30: 772-780.
38. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 2003; 52: 696-704.
39. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, LargetB; Liu L, Suchard MA, Huelsenbeck JP : MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012; 61: 539-542.
40. Gouy M, Guindon S, Gascuel O: SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Bio Evol. 2010; 27: 221-224.
41. Yang Z: PAML 4: phylogenetic analysis by maximum likelihood. Mol Bio Evol. 2007; 24(8): 1586-1591.
42. Dash S, Cannon EKS, Kalberer SR, Farmer AD, Cannon SB: PeanutBase and other bioinformatic resources for peanut (Chapter 8). In peanuts genetics, processing, and utilization, edited by Stalker HT and WilsonvRF. AOCS Press; 2016. P. 241–252. ISBN 9781630670382.
43. Clevenger J, Chu Y, Scheffler B, Ozias-Akins P: A developmental transcriptome map for allotetraploid *Arachis hypogaea*. Front Plant Sci. 2016;7:1446.
44. Chang S, Puryear, JJC: A simple and effiffifficient method for isolating RNA from pine trees. Plant Mol Biol Rep.1993;11:113-116.
45. Xia H, Zhao C, Hou L, Li A, Zhao S, Bi Y, An J, Zhao Y, Wan S, Wang X: Transcriptome profiling of peanut gynophores revealed global reprogramming of gene expression during early pod development in darkness. BMC Genomics. 2013; 14:517.
46. Langfelder P, Horvath S: WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559.

Supplementary Material

Additional file 1: Primers used in this study.

Additional file 2: The number of bases between the cysteins in AdLTPs. Values in red are greater than those previously reported by Edstam et al. for other species [2].

Additional file 3: Transcription factor binging to promoters of *AdLTP* genes associated with resistance to nematode infection.

Additional file 4: Genes that may interact with *AdLTP* genes involved in resistance to nematode infection.

Figures

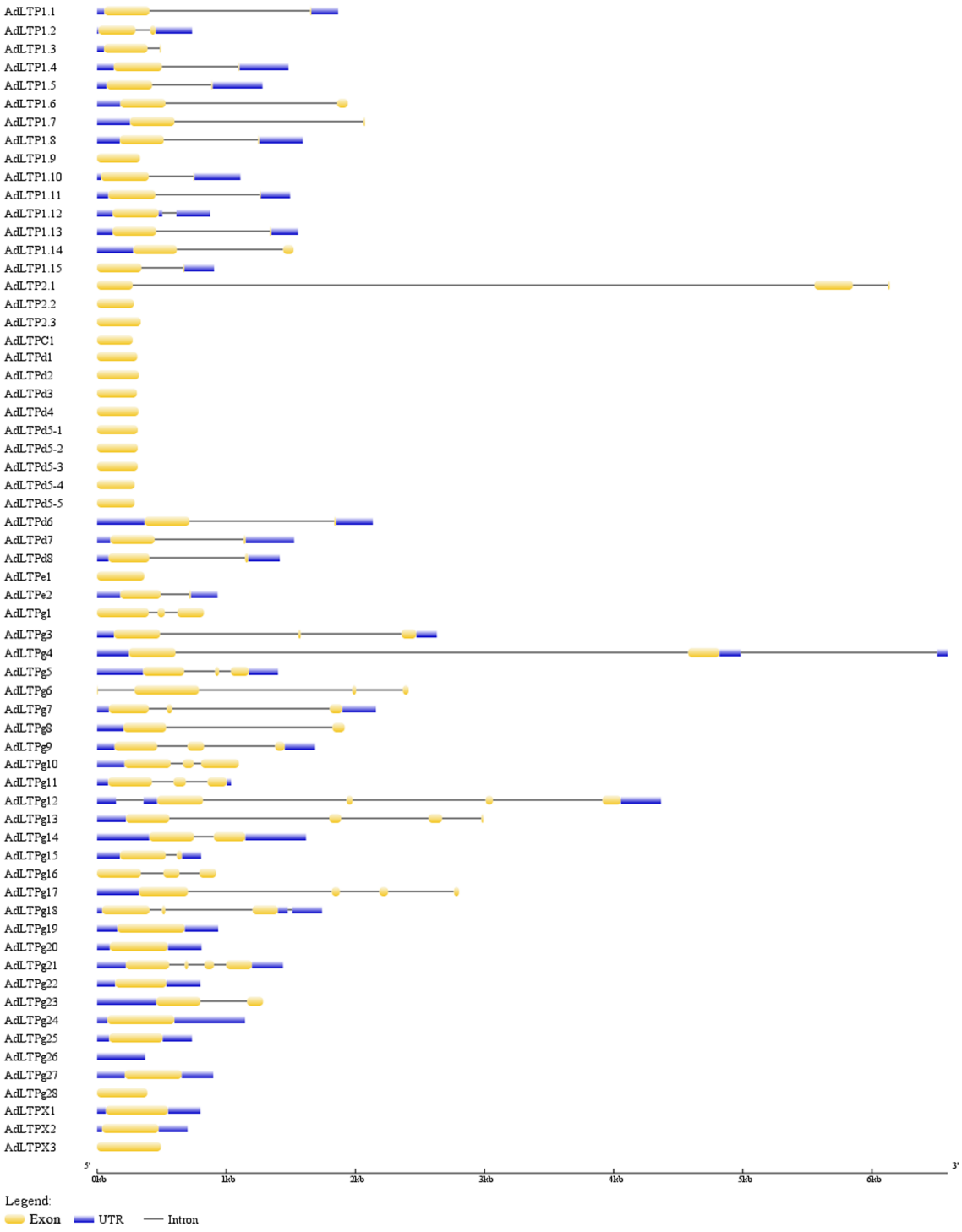

Figure 1

Analysis of exon–intron structure of AdLTP genes. UTRs, exons, and introns are represented by dark blue, yellow, and black lines, respectively.

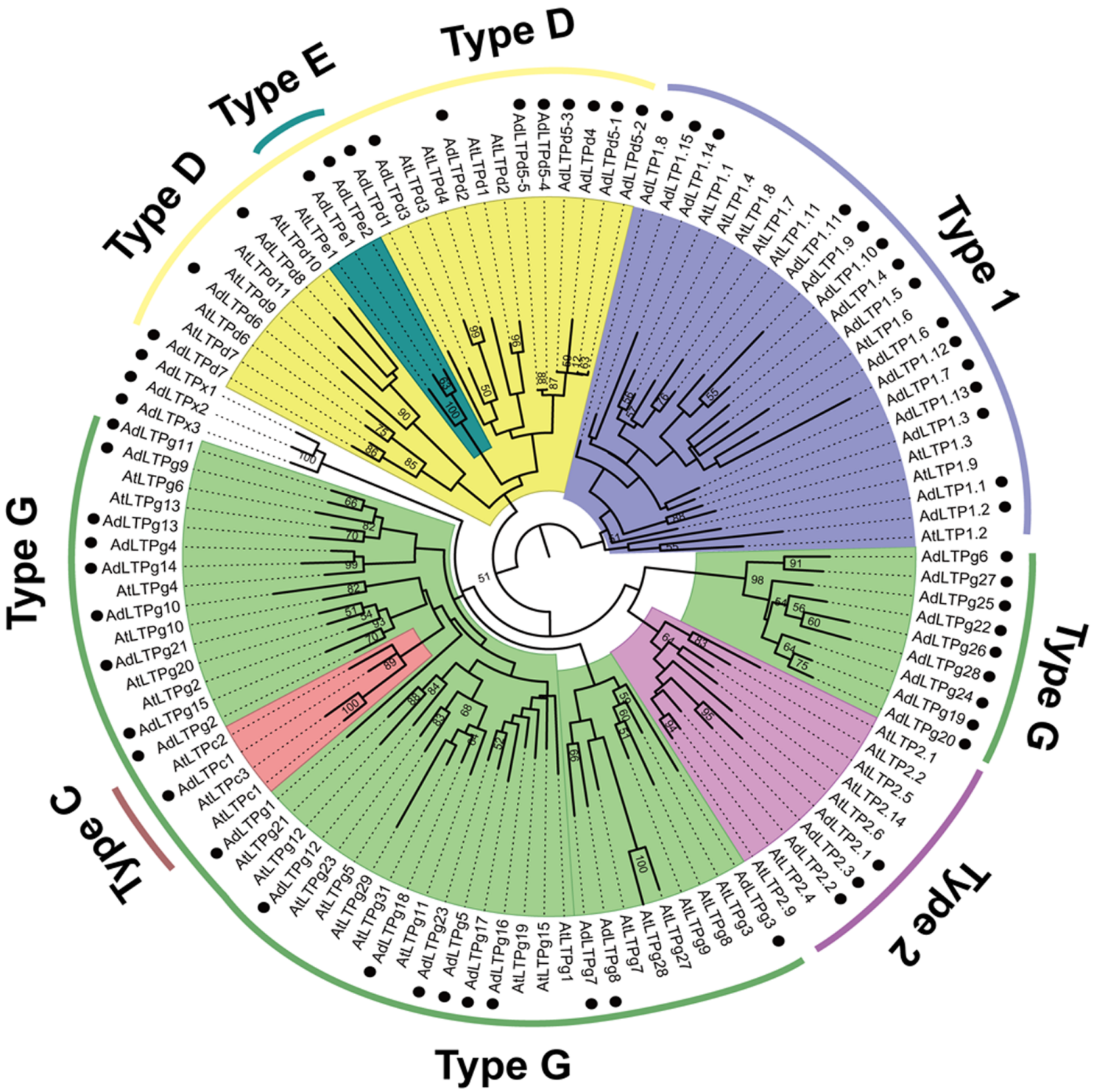

Figure 2

Phylogenetic tree of nsLTPs from A. duranensis and Arabidopsis thaliana. The nsLTPs could be classified into six subfamilies: Types 1, 2, C, D, E, and G. Black dots indicate the AdLTPs.

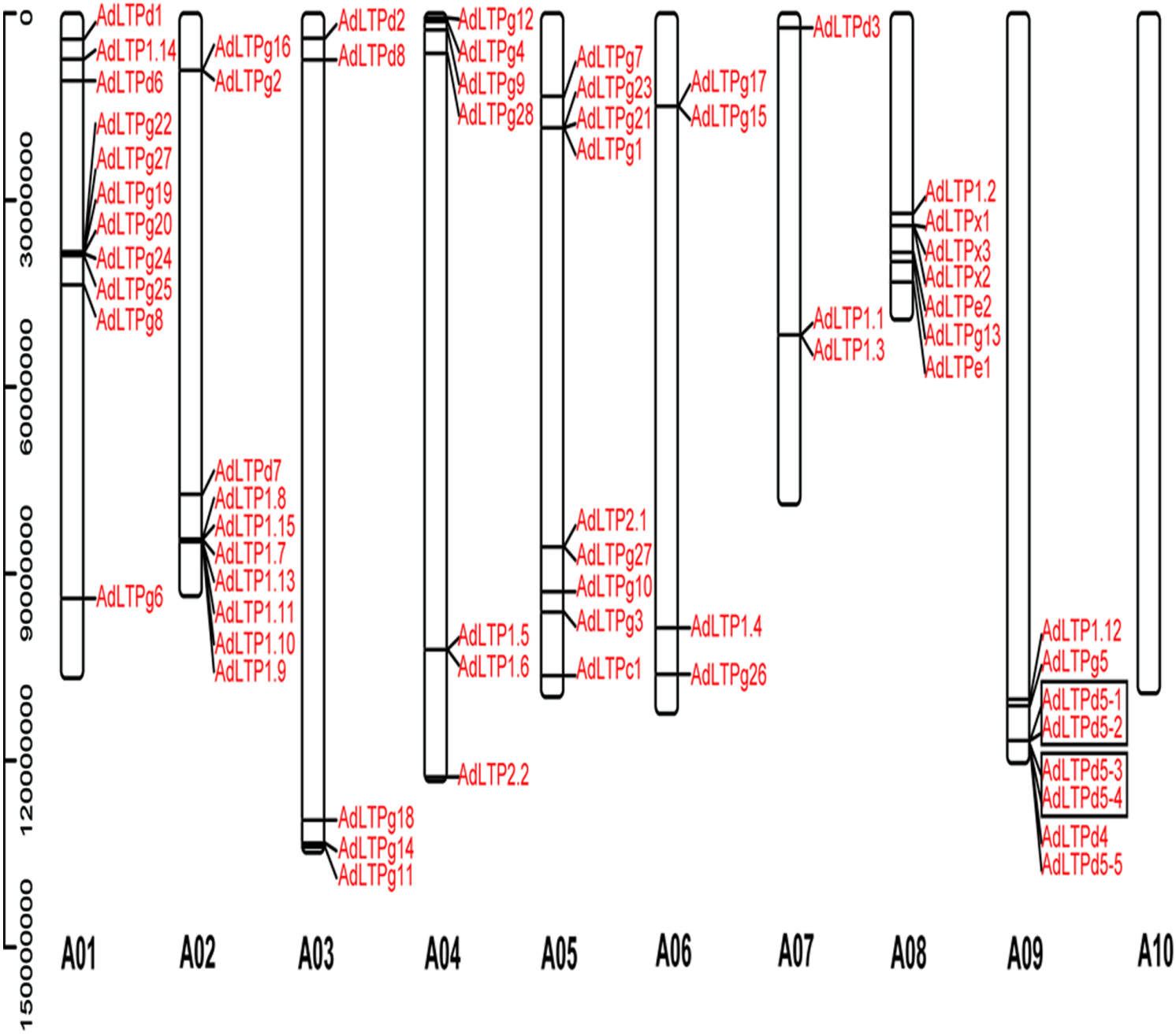

Figure 3

Chromosomal location of AdLTP genes. Chromosome numbers are indicated at the bottom. The location of each AdLTP gene is indicated by a horizontal line. The black boxes indicate the tandem duplication genes.

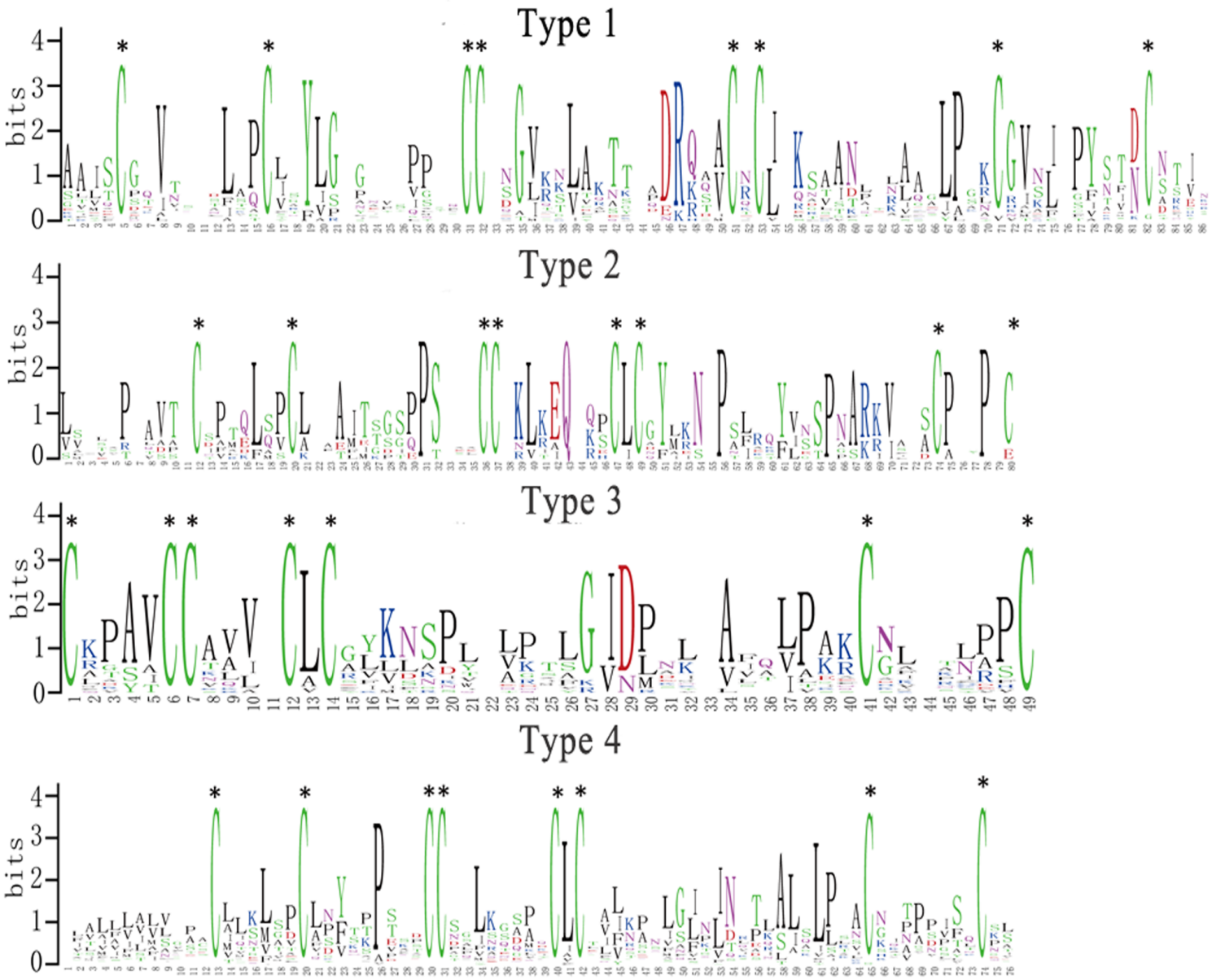

Figure 4

Domains with eight conserved cysteine residues (8CM) of Types 1, 2, D, and G in A. duranensis. The conservation of AdLTPs was detected by weblogo online software, and highly variable parts of the alignment were determined and manually excluded in SeaView v4. All detected nsLTPs contained a conserved C-Xn-C-Xn-CC-Xn-CXC-Xn-C-Xn-C domain (where C is cysteine, and X is any other amino acid). The size of the letter indicates the degree of conservation.

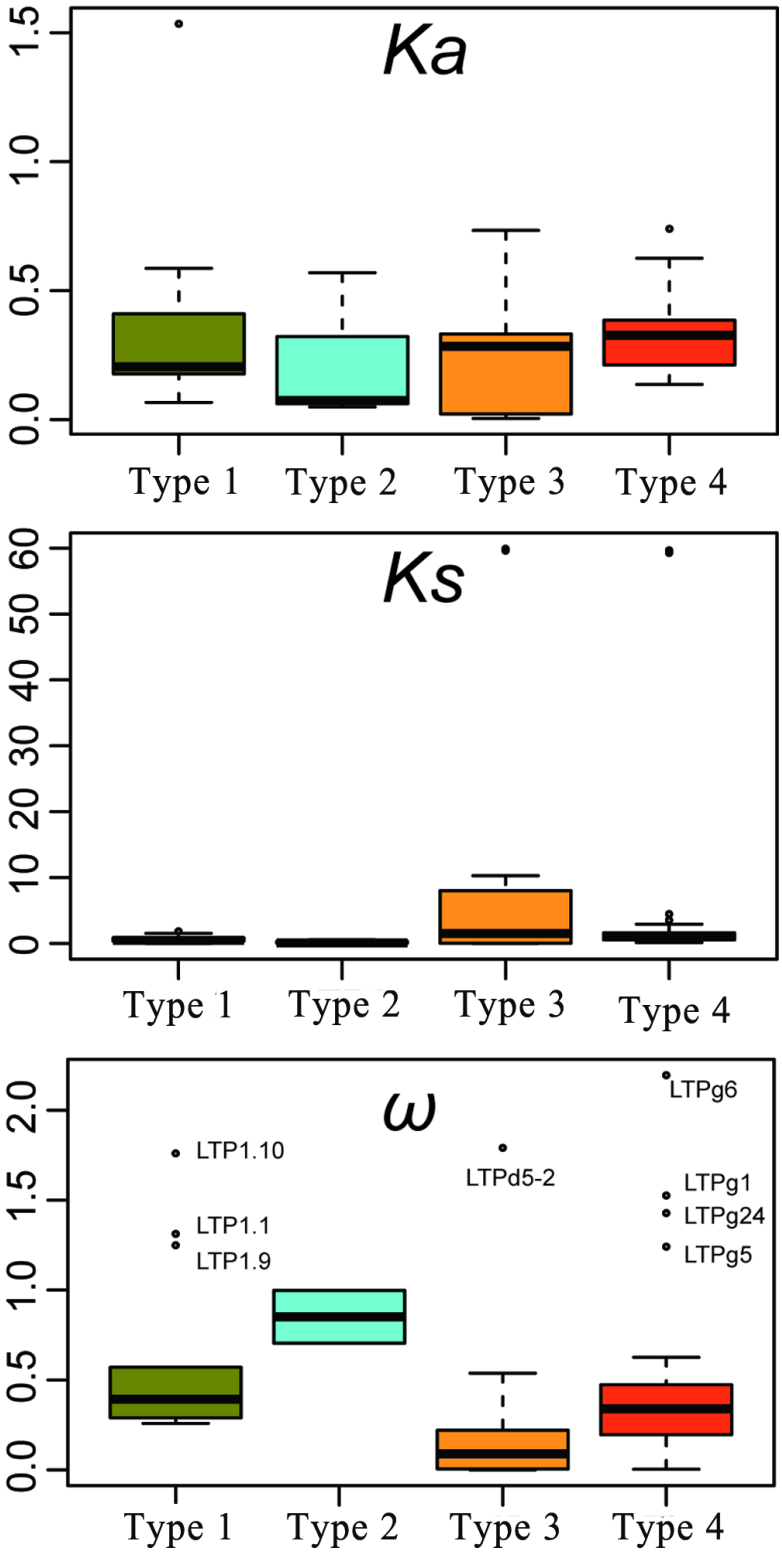

Figure 5

The nonsynonymous substitution rate (Ka), synonymous substitution rate (Ks), and nonsynonymous/synonymous substitution (ω) rate of subfamilies: Types 1, 2, D, and G in A. duranensis. Eight genes (LTP1.1, LTP1.9, LTP1.10, LTPd5-2, LTPg1, LTPg5, LTPg6, and LTPg24) were under positive selection, indicating that certain AdLTPs might have different functions.

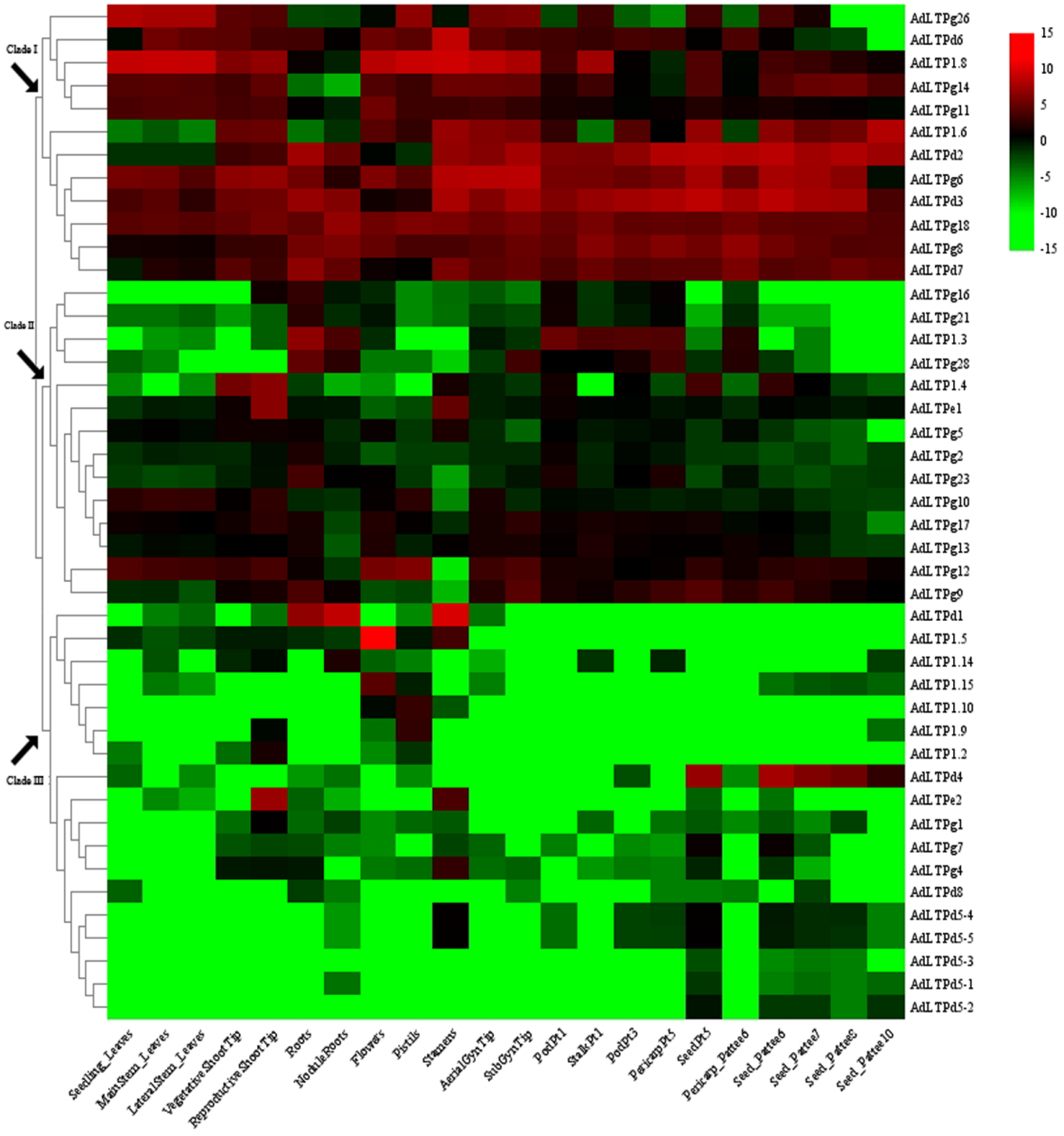

Figure 6

The expression of AdLTP genes in 22 A. duranensis tissues. Seedling_Leaves, seedling leaf 10 d post emergence; MainStem_Leaves, main stem leaf; LateralStem_Leaves, lateral leaf; VegetativeShootTip, vegetative shoot tip from the main stem; ReproductiveShootTip, reproductive shoot tip from the first lateral leaf; Roots, 10 d roots; NoduleRoots, 25 d nodules; Flowers, perianth; Pistils, gynoecium; Stamens, androecium; AerialGynTip, aerial gynophore tip; SubGynTip, subterranean gynophore tip; PodPt1, Pattee stage 1 pod; StalkPt1, Pattee stage 1 stalk; PodPt3, Pattee stage 3 pod; Pericarp_Pattee5, Pattee stage 5 pericarp; Seed_Pattee5, Pattee stage 5 seed; Pericarp_Pattee6, Pattee stage 6 pericarp; Seed_Pattee6, Pattee stage 6 seed; Seed_Pattee7, Pattee stage 7 seed; Seed_Pattee8, Pattee stage 8 seed; and Seed_Pattee10, Pattee stage 10 seed.

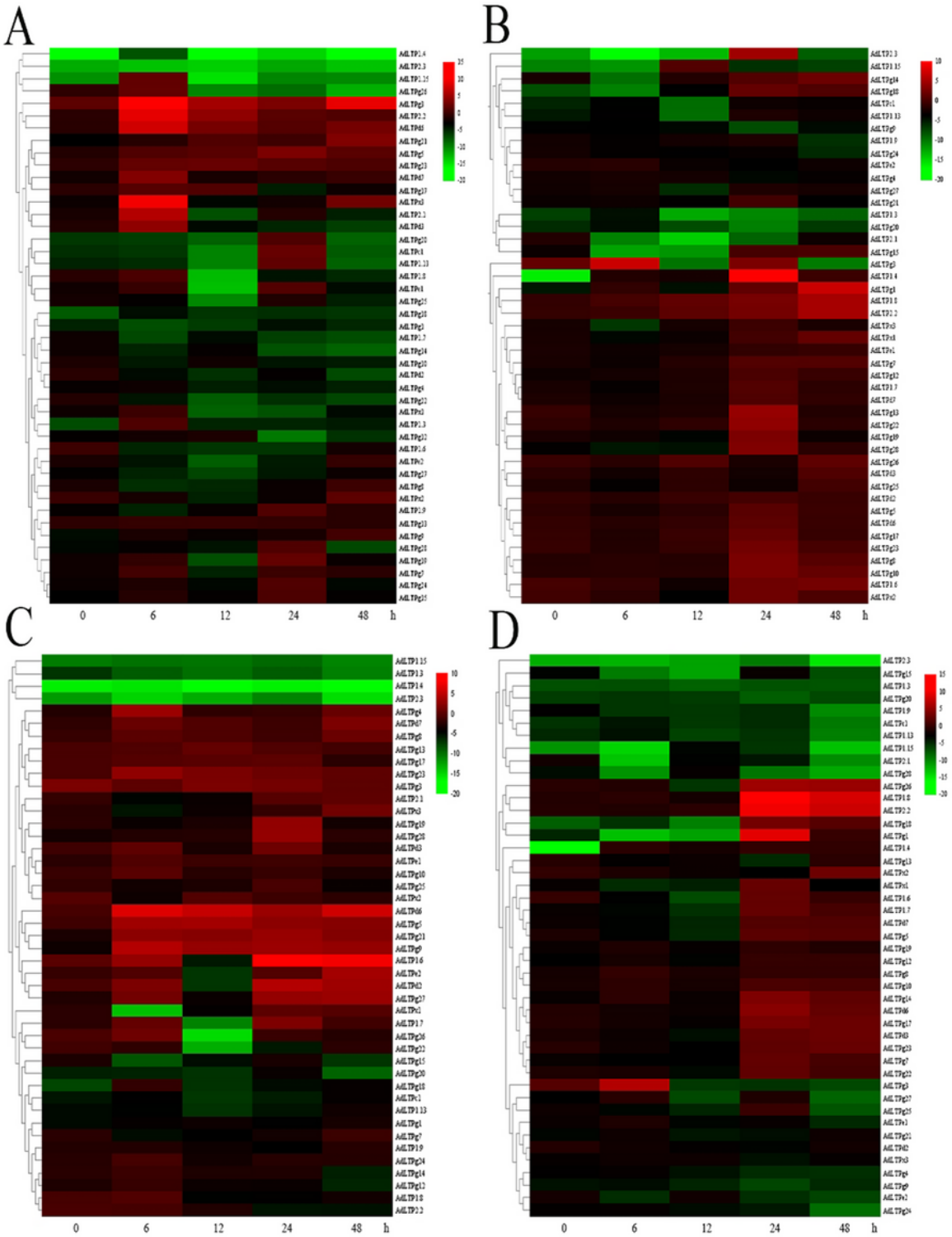

Figure 7

The expression of AdLTPs in A. duranensis seedlings exposed to the following stresses: PEG (A), low temperature (B), NaCl (C), and ABA (D).

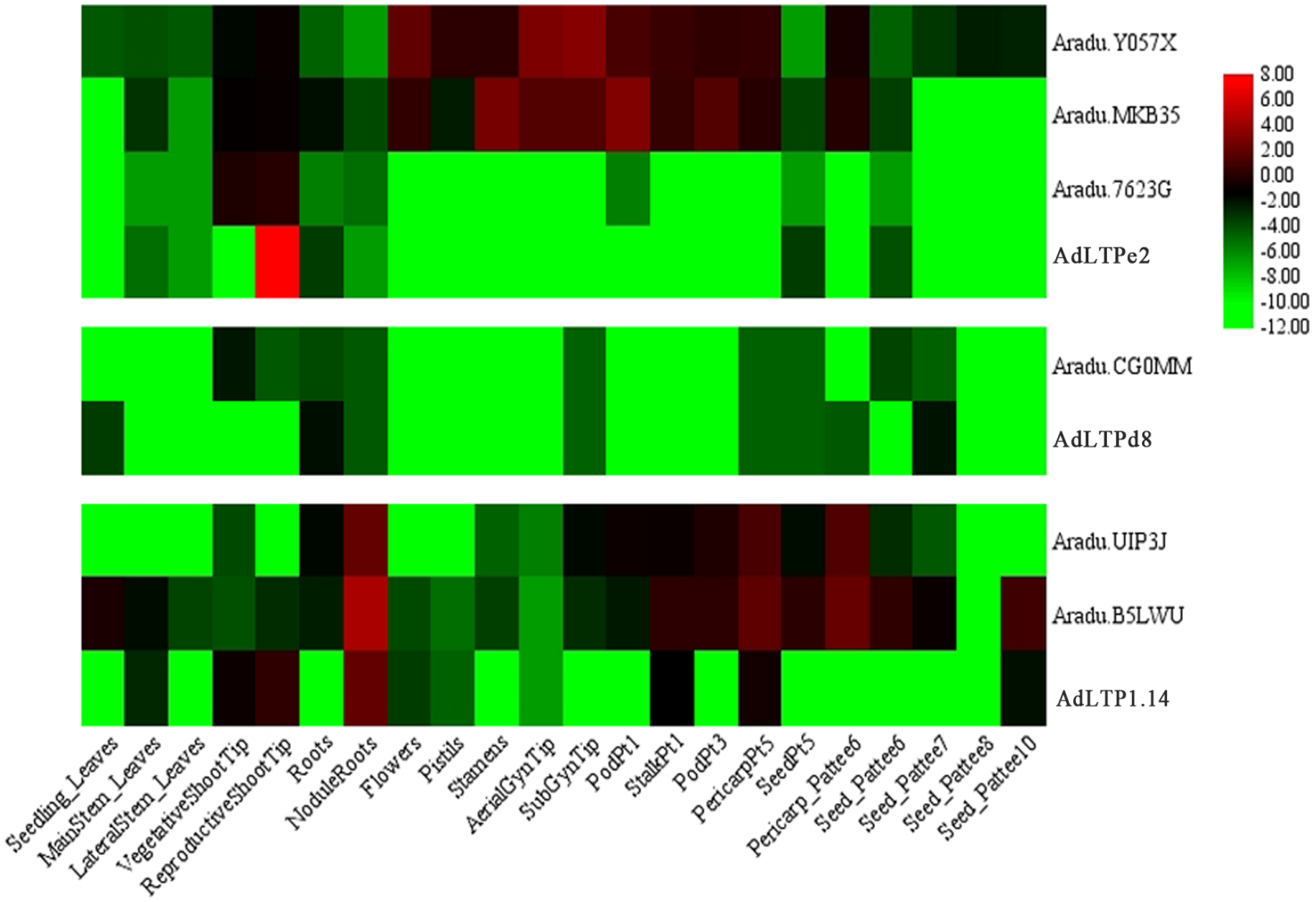

Figure 8

The similar expression patterns between AdLTP genes and other genes co-expressed during resistance to nematode in response to nematode infection. Red indicates upregulation, and green indicates downregulation.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile4.docx
- Additionalfile2.docx
- Additionalfile1.docx
- Additionalfile3.docx